

Dynamic Defocus and Occlusion Compensation of Projected Imagery by Model-Based Optimal Projector Selection in Multi-projection Environment

Momoyo Nagase · Daisuke Iwai · Kosuke Sato

Received: date / Accepted: date

Abstract This paper presents a novel model-based approach of dynamic defocus and occlusion compensation method in a multi-projection environment. Conventional defocus compensation research applies appearance-based method which needs a PSF (point spread function) calibration when either position or orientation of an object to be projected is changed, thus cannot be applied to interactive applications in which the object dynamically moves. On the other hand, we propose a model-based method in which PSF and geometric calibrations are required only once in advance, and projector's PSF is computed online based on geometric relationship between the projector and the object without any additional calibrations. We propose to distinguish the oblique blur (loss of high-spatial-frequency components according to the incidence angle of the projection light) from the defocus blur and to introduce it to the PSF computation. For each part of the object surfaces, we select an optimal projector which preserves the largest amount of high-spatial-frequency components of the original image to realize defocus-free projection. The geometric relationship can also be used to eliminate the cast shadows of the projection images in multi-projection environment. Our method is particularly useful in the interactive systems because the movement of the object (consequently geometric relationship between each projector and the object) is usually measured by an attached tracking sensor. This paper describes details about the proposed approach and a prototype implementation. We performed two proof-

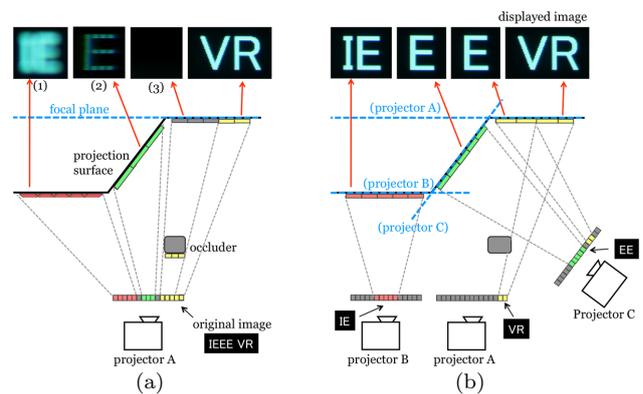


Fig. 1 Inherent optical problems of video projectors and their multi-projector solution. (a) In a single-projector setup, the image quality of the displayed content is decreased by (1) defocus blur, (2) oblique blur, and (3) cast shadows. (b) The proposed approach applies multiple projectors to solve the above issues to improve the image quality.

of-concept experiments to show the feasibility of our approach.

Keywords Projection-based mixed reality · Multi-projection environment · Defocus compensation · Shadow removal · PSF computation

1 Introduction and Motivation

Projection-based mixed reality (MR) technologies are widely applied to interactive systems in which users interact with real objects that are visually augmented by projection images, such as a real object painting system (Bandyopadhyay et al, 2001) and an interactive book (Gupta and Jaynes, 2006). In addition, recent reduction in size and price of video projectors could lead to many other interactive applications. One important advantage of projection-based MR over other types of

MR technologies is that users do not have to wear any special and annoying devices such as a head-mounted display (Bimber and Raskar, 2005). However, video projectors that emit images through an optical lens inherit three problems (Fig. 1). First, **defocus blur** degrades high-spatial-frequency components of the original image. Second, the high-spatial-frequency component also attenuates when the incidence angle of the projection light is not perpendicular to the surface. In this paper, we refer to this effect as **oblique blur**. The last problem is the **cast shadow** of the projection light.

Apertures of video projectors are designed to be wide, such as F1.4, to ensure a large amount of light throughput. Such large stops results in a narrow depth-of-field. Once the projection surface moves along the optical axis even for a short distance, the displayed contents are not always in focus. Oblique blur makes the projected pixel spread anisotropically on the surface. In interactive systems, these defocus and oblique blurs occur frequently because projection objects are dynamically moved by users' actions. Furthermore, when projection light is blocked by an object that exists between the projector and the surface, a shadow casts on the surface. Therefore, users should not enter any objects between the projector and the surface. However, this constraint is generally unfavorable for interactive systems.

In this study, we propose a novel model-based approach of dynamic defocus and occlusion compensation method in a multi-projection environment, which simultaneously takes into account the oblique blur. We assume that geometric information (initial relative rigid-body transforms among the projectors and the object, and the shape of the object) is calibrated in advance, and that 6 degrees-of-freedom (DOF) motion of the moving object is measured online. Note that this is a general assumption of most interactive systems (Bandyopadhyay et al, 2001; Low et al, 2001; Kondo and Kijima, 2002; Gupta and Jaynes, 2006; Park et al, 2006; Kondo et al, 2008). Once the geometric calibration is performed, the proposed method provides online defocus- and occlusion-free projection without any additional calibration even when the object moves.

For each point of the object surface, the proposed system calculates a point spread function (PSF) for each projector based on the relative rigid-body transform of the projector to the point. In the PSF calculation, the system simultaneously deals with the defocus and the oblique blurs. Then, it analyzes the PSFs and selects the optimal one that preserves the largest amount of high-spatial-frequency components of the original image. Consequently, high spatial resolution images are displayed from the selected optimal projector on the

point. Furthermore, a visibility test that checks whether the point is visible from each projector is also operated with the geometric information. The result of the visibility test is used to achieve occlusion-free projection.

Specifically, we make the following contributions:

- dynamic defocus and occlusion compensation of projected imagery in which no online calibration is needed
- a novel model-based PSF computation based on geometric relationship between target surface and projector
- the introduction of the oblique blur to improve the PSF computation

The remainder of the paper is organized as follows. The subsequent section briefly describes related studies. Section 3 presents the detailed principles of the proposed dynamic defocus and occlusion compensation method. Section 4 describes the detailed implementation of the proposed system and shows proof-of-concept experiments. In Section 5, we briefly discuss advantages and limitations of the proposed method and directions for future work, and we conclude the paper in Section 6.

2 Related Work

Two research areas are closely related to our approach: defocus compensation and occlusion compensation of projected imagery.

2.1 Defocus Compensation

Several approaches for deblurring unfocused projections using a single or multiple projectors have been proposed in the field of projector-camera systems (Bimber et al, 2008).

Single projector approaches digitally sharpen the original image before projection, so that an optically defocused projection appears close to the original (i.e., un-blurred) image. Defocus blur of a projection image results from a convolution of a PSF and the original image. If the PSF of the projector on the object surface is correctly estimated, a defocus-free image can be displayed by the Wiener filter calculation in the frequency domain (Brown et al, 2006). Zhang and Nayar formulated the image correction as a constrained optimization problem in the spatial domain using a convolution matrix (Zhang and Nayar, 2006). They estimated the PSF by projecting pattern images such as 2D array of dots or crosses and capturing the displayed images in advance. In addition to the deconvolution process, Grosse and Bimber proposed to increase the optical depth-of-field of a projector by applying a coded

aperture (Grosse and Bimber, 2008). However, these approaches cannot be applied to interactive systems because they need to project pattern images when the objects move, and they consequently disturb users' natural interactions. To solve this problem, Oyamada and Saito estimated the PSF just by observing the displayed content so that they did not need to project special pattern images (Oyamada and Saito, 2008). However, the defocus blur cannot always be compensated through digital processing alone. Because the PSF of a normal circular aperture contains many zeros in the frequency domain, high-spatial-frequency components are lost in out-of-focus regions. Consequently ringing artifacts are often led when direct inverse methods such as Wiener deconvolution are applied to compensate the optically low-pass filtered projected light. Furthermore, a single projector approach cannot realize occlusion-free projection.

On the other hand, Bimber and Emmerling applied multiple projectors, each of which has the focal plane at a unique position, to realize multifocal, or defocus-free, projection (Bimber and Emmerling, 2006). For each point of the object surface, they selected an optimal projector that could display the finest image on the point. Their multi-projector approach does not need to perform the deconvolution as do the above single projector approaches. However, it also needs to project special pattern images on the surface to estimate the PSF from every projector when the object moves and thus is not suitable for interactive systems.

2.2 Occlusion Compensation

Occlusion-free projection is achieved using multiple projectors. Most of the previous methods applied appearance-based approaches where cameras were used to detect shadows on the projection screen (Sukthankar et al, 2001; Cham et al, 2003; Jaynes et al, 2004). On the other hand, a model-based approach was also proposed by Raskar et al. where geometric information of the projectors and the object surfaces (relative rigid-body transforms among the projectors and the surfaces, and shapes of the surfaces) is assumed to be known (Raskar et al, 2001). Then, a visibility test is performed for each point of the object surfaces from each projector. Consequently, occlusion (i.e., cast shadow) can be removed because images are displayed from a projector that is visible from the point. When the position of the object is measured online, cast shadows can be removed even for the dynamic objects, as proposed in (Audet and Cooperstock, 2007). These studies, however, did not merge the shadow removal techniques with defocus compensation.

2.3 Our Approach

We apply a multi-projector approach to achieve defocus compensation of projection images where any digital deconvolution processing is not needed, as in (Bimber and Emmerling, 2006). For each point of the object surface, our approach selects the optimal projector that can project the finest image on the point, and the selected projector displays images there. Furthermore we apply an occlusion compensation approach and merge it with the defocus compensation of projected imagery.

Unlike the previous work (Bimber and Emmerling, 2006) which is an appearance-based approach where PSFs are measured by a camera, our method is a model-based approach where PSFs are computed by geometric information of the projectors and the objects. Our method does not need to project special calibration pattern images on the object surface for the PSF computation, therefore it can be applied to interactive systems in which the object surface frequently moves. Furthermore, our method computes the shape of the PSF by considering that the frequency property of the PSF depends on the surface orientation while the above previous deblurring techniques did not.

3 Defocus and Occlusion Compensation Method

This section describes our dynamic defocus and occlusion compensation approach in a multi-projector environment. For each point of the object surface, our approach selects an optimal device that projects the finest images on the point by analyzing PSFs of all the projectors. Therefore, selection of the optimal projector is synonymous with defocus compensation in our context.

In Section 3.1, we explain how to select the optimal projector that can display the finest images on the object surface. In Section 3.2, we describe a PSF model of projected pixels and explain how to apply the model to the proposed projector selection method. In Section 3.3, we show how to integrate an occlusion compensation method into our defocus compensation approach. Section 3.4 describes a feathering technique for seamless merging of images from multiple projectors.

3.1 Spatial Frequency Analysis of PSF and Optimal Projector Selection

A projected image is defocused on the object surface according to its PSF. The process described below is for a point of the object surface. The same process is also applied to the other points. Suppose that the PSF

of a projector p is represented as $PSF_{II}(p)$, and that the original image is represented as i . When i is projected onto the point, the displayed result o is calculated by the following function:

$$o = i * PSF_{II}(p), \quad (1)$$

where $*$ represents a convolution operation, and II means that the PSF is a two-dimensional (2D) function.

A PSF of defocus blur can be considered as a low-pass filter because it attenuates high-spatial-frequency components of the original image. Thus, we focus on its frequency property, particularly the cutoff frequency, and define the optimal projector as one whose PSF's cutoff frequency is the highest.

A PSF is transformed into the optical transfer function (OTF) in the frequency domain. Suppose that the Fourier transforms of $PSF_{II}(p)$, the original image i , and the displayed image o are denoted by $OTF_{II}(p)$, I , and O respectively. The convolution of Eq. (1) is transformed into a simple multiplication in the frequency domain:

$$O = I \cdot OTF_{II}(p). \quad (2)$$

Although the OTF contains both the magnitude and phase components, we consider only the magnitude that is known as the modulation transfer function (MTF). The cutoff frequency, which is a term used in physics and electrical engineering, means a boundary in a system's frequency response at which energy flowing through the system begins to be attenuated rather than passing through. In the case of a low-pass filter, more than 50% of the signal is attenuated when the frequency is higher than the cutoff frequency. This means that we can know that at least 50% of the contrast of the original is preserved in a frequency band that is lower than the cutoff frequency by analyzing the MTF.

Suppose that the cutoff frequency of a projector p is represented as $f_c(p)$. We define the optimal projector as one whose cutoff frequency is the highest. Thus, the optimal projector p_{opt} is selected as:

$$p_{opt} = \arg \max_p f_c(p). \quad (3)$$

The incidence angle of projection light is not always perpendicular to the object surface. When it is not perpendicular, the PSF becomes anisotropic (i.e., the oblique blur). Thus, the one-dimensional (1D) projection of the PSF changes according to the direction of the projection θ . The MTF of the 1D projection of the PSF and its cutoff frequency also change according to θ . The 1D projection of the PSF, its MTF, and its cutoff frequency are represented as $PSF_I(p, \theta)$, $MTF_I(p, \theta)$,

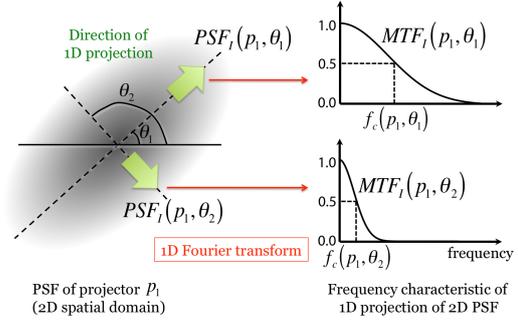


Fig. 2 Anisotropic frequency characteristic of a PSF of a projected pixel

and $f_c(p, \theta)$, respectively. We search for a 1D projection of the PSF whose cutoff frequency is the lowest, and define it as the cutoff frequency of the projector p as:

$$f_c(p) = \min_{\theta} f_c(p, \theta). \quad (4)$$

Figure 2 illustrates an example of an anisotropic frequency characteristic of a PSF from a projector p_1 . Two MTFs, $MTF_I(p_1, \theta_1)$ and $MTF_I(p_1, \theta_2)$, are shown in the right of the figure. They are Fourier transforms of the 1D projections of the anisotropic PSF with the different projection directions θ_1 and θ_2 . As shown in the figure, the cutoff frequency of $MTF_I(p_1, \theta_1)$ is higher than the other; thus, $f_c(p_1, \theta_1) > f_c(p_1, \theta_2)$. In this way, we compare the cutoff frequencies for all the projection directions, determine the lowest one, and designate it as the cutoff frequency of the projector p_1 (i.e., $f_c(p_1)$).

The basic principle of our approach is summarized as follows:

For each point of the surfaces,

1. Compute the lowest cutoff frequency of each projector (Eq. (4))
2. Find the optimal projector that gives the highest cutoff frequency among the computed lowest cutoff frequencies (Eq. (3))

3.2 PSF Model

We introduce a PSF model based on lens optics, and then apply it to our optimal projector selection approach described in Section 3.1.

3.2.1 Geometry of Defocus Blur and Oblique Blur

In a projector's optical system, the focal plane is perpendicular to the lens axis situated at the focal-length distance from the lens. Suppose that an infinitesimal point light source on the projector plane is projected through the lens onto a surface perpendicular to the

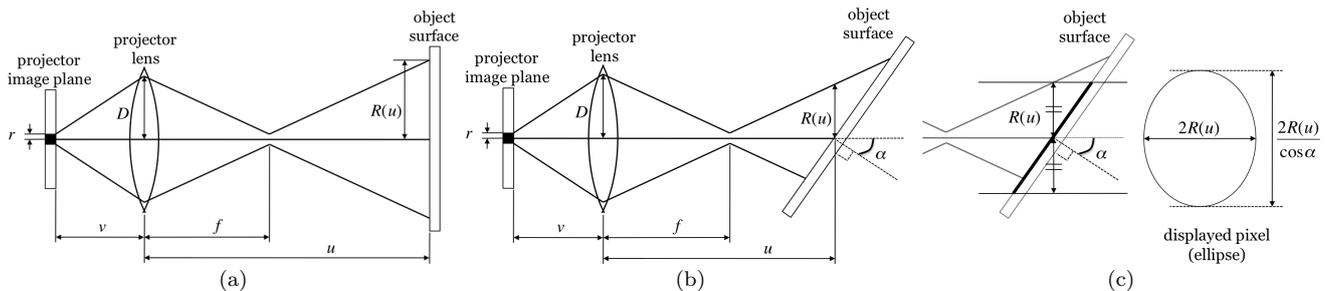


Fig. 3 Lens optics of a projector. (a) A circle whose radius is r (in the projector plane) is projected on a surface perpendicular to the lens axis. (b) The circle is projected on the surface by a different incidence angle. (c) PSF ellipse.

lens axis. The radius $R(u)$ of the blur circle on the projection surface is calculated as:

$$R(u) = |D(\frac{u}{f} - 1)|, \quad (5)$$

where f is the projector focal length, D is the radius of the projector lens, and u is the distance of the surface from the lens.

In practice, projectors cannot produce infinitesimally small light sources. If instead we project a circle whose radius is r (in the projector plane), the radius $R(u)$ of the blurred circle on the surface is calculated as:

$$R(u) = |D(\frac{u}{f} - 1)| + \frac{ru}{v}, \quad (6)$$

where v is the distance of the projector plane from the lens (Fig. 3(a)).

In the case that the incidence angle of the projection light is not perpendicular to the surface as shown in Fig. 3(b), the projected pixel forms an ellipse (the intersection of a cone with a plane). Suppose that the distance of the surface from the projector lens is sufficiently longer than the radius of the lens; we approximate the cone as a cylinder. Thus, the projected pixel on the surface forms an ellipse that is the intersection of the cylinder and the surface (Fig. 3(c)). The radius of the cylinder's bottom circle is $R(u)$. Consequently, the length of the semi-major axis e_a and one of the semi-minor axis e_b are:

$$e_a = \frac{R(u)}{\cos \alpha}, \quad e_b = R(u), \quad (7)$$

where α is the incidence angle. We refer to this ellipse as the **PSF ellipse** in this paper.

3.2.2 Intensity Roll-off

In general, the intensity of a blurred pixel decreases from the center to the edge of the PSF ellipse. While the PSF depends on the lens system, it can be reasonably modeled as a 2D Gaussian distribution (Brown et al, 2006). We apply a PSF model that approximates the

PSF as an anisotropic 2D Gaussian distribution whose level set forms the PSF ellipse. 2D Gaussian distribution is defined by two variances: we refer to the greater one as the **major variance** and the other as the **minor variance** because they correspond to the major axis and the minor axis of the ellipse, respectively. Suppose that the major and the minor variances of a projector p are represented as $\sigma_a(p)$ and $\sigma_b(p)$ respectively, the PSF is modeled as:

$$PSF_{II}(x, y|p) = l(p) \exp\left\{-\frac{1}{2}\left(\frac{x^2}{\sigma_a^2(p)} + \frac{y^2}{\sigma_b^2(p)}\right)\right\}, \quad (8)$$

where $l(p)$ represents the intensity value at the center, and $e_a(p)$ and $e_b(p)$ represent the length of the semi-major and the semi-minor axes of the PSF ellipse of the projector p respectively. Note that $\sigma_a(p)$ is greater than $\sigma_b(p)$, and $\sigma_a(p)$ and $\sigma_b(p)$ are proportional to $e_a(p)$ and $e_b(p)$, respectively:

$$\sigma_a(p) = k(p)e_a(p), \quad \sigma_b(p) = k(p)e_b(p). \quad (9)$$

where $k(p)$ is a proportionality coefficient.

3.2.3 Optimal Projector Selection

The normalized MTF of the PSF introduced in Eq. (8) is:

$$MTF_{II}(w_x, w_y|p) = \exp\left\{-\frac{1}{2}\left(\frac{w_x^2}{\sigma_a^2(p)} + \frac{w_y^2}{\sigma_b^2(p)}\right)\right\}, \quad (10)$$

where $w_x = 2\pi f_x$ and $w_y = 2\pi f_y$.

We apply this MTF to the optimal projector selection proposed in Section 3.1. According to the central slice theorem of computed tomography, a Fourier transform of the 1D projection of a 2D function is same as the 1D slice of the 2D Fourier transform of the function (the slice crosses the center and its direction is perpendicular to the direction of the projection). From Eq. (10), the MTF is also a 2D Gaussian distribution whose level set forms an ellipse. The cutoff frequency of the 1D slice of the MTF becomes the lowest when

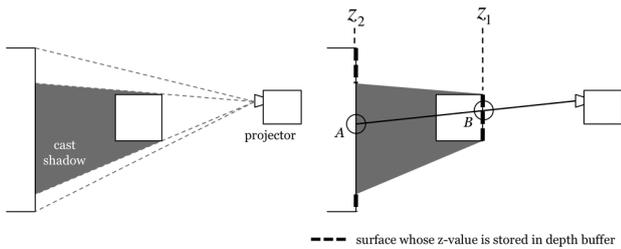


Fig. 4 Shadow area detection. (left) A sample projection scene. (right) A point A is recognized as a shadow area and a point B is not.

the direction of the slice is along the minor axis of the ellipse. The 1D slice is represented as:

$$MTF_I(w|p) = \exp\left(-\frac{w^2}{2\sigma_a^2(p)}\right). \quad (11)$$

Equation (11) indicates that the lowest cutoff frequency increases as the major variance $\sigma_a(p)$ decreases.

According to the basic principle described in Section 3.1, we at first compute the major variance that is related to the lowest cutoff frequency of each projector p . Then, we find the optimal projector whose computed major variance is the minimum among all the projectors, since the minimum major variance gives the highest cutoff frequency. So now, Eq. (3) is rewritten as:

$$p_{opt} = \arg \min_p \sigma_a(p). \quad (12)$$

3.3 Occlusion Compensation

When multiple projectors exist, we can remove the cast shadows by displaying images on the shadow areas from projectors that are visible from the areas. This is accomplished when the geometric information (rigid-body transforms of the projectors and shape of the target object) is known. Because our approach assumes that there are multiple projectors in the system and the geometric information is calibrated in advance, occlusion-free projection can be performed.

For each point of the object surface, our system determines whether each projector casts a shadow on the point by a visibility test (Fig. 4). First, our system performs off-screen rendering of the projection scene from each projector viewpoint and stores the generated depth buffer. Next, it computes the distance of each point of the surface from each projector and compares it with the corresponding depth value stored in the depth buffer. If these two values are not the same, the point of the surface is determined as a shadow area from the projector (Fig. 4).

We integrate a shadow removal process into the optimal projector selection method introduced in Sections

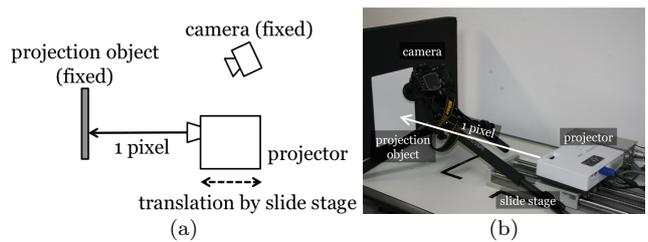


Fig. 5 PSF calibration. (a) Diagram. (b) Actual configuration.

3.1 and 3.2. For each point of the object surface, our approach selects the optimal projector according to Eq. (12) among the projectors that do not cast shadows on the point. Consequently, defocus and occlusion compensation can be simultaneously realized.

3.4 Feathering of Multiple Projector Contribution

A feathering technique is used to merge seamlessly images from multiple projectors. It is needed because of the lack of color equivalence among the projectors due to manufacturing processes and temperature color drift during their use, and because of our desire to minimize the sensitivity to small errors in the estimated geometric calibration parameters. Thus, we apply one of the feathering techniques that have been already proposed by several researchers (Raskar et al, 2001; Bimber and Emmerling, 2006; Majumder and Brown, 2007).

4 Experiment

We conducted three experiments. The projector used in the following experiments was EPSON EMP-1710 (resolution: 1024×768 pixels, aperture: F1.6). A PC (CPU: 3.4 GHz, RAM: 2.0 GB) was used to control the devices. First, we measured the actual variances of the PSFs by changing the distance of the projector from the projection surface. Second, we validated the feasibility of our dynamic defocus compensation approach by comparing it with another one that does not explicitly deal with the oblique blur. Third, we performed another proof-of-concept experiment to confirm whether defocus and occlusion compensation was simultaneously achieved for a dynamic scene.

4.1 PSF Calibration

The major variance σ_a of a PSF is needed in the optimal projector selection, as shown in Eq. (12). Because there are several unknown parameters in Eq. (6) such as D , f , r , and v , we measured actual PSFs that will be used in the following experiment.

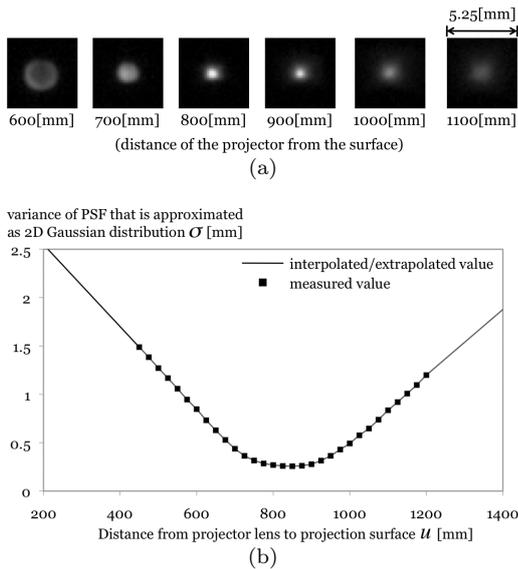


Fig. 6 Calibration result. (a) Captured pixel. (b) Variance of the captured pixel that is approximated as 2D Gaussian distribution.

We placed a projector on a slide stage that moved along its optical axis, a planar surface that was perpendicular to the axis, and a camera that was directed at the surface in a dark room (Fig. 5). The projector projected a white pixel on the surface from different distances, and the camera captured the projected pixel. The distance of the projector lens from the surface was varied from 450 mm to 1200 mm at 25 mm intervals. Several captured results are shown in Fig. 6(a). These captured pixels were approximated as 2D Gaussian distributions, and their variances are plotted in Fig. 6(b).

The measured variances were linearly interpolated to fill the missing values within the measured range. The missing values outside the range were linearly extrapolated by the least-squares method because the variance of a PSF linearly decreases until the focal plane and then increases (Fig. 6(b)). From the result, we defined the focal distance of the projector as 850 mm, where the variance is minimum.

The obtained variance was the minor variance of the PSF, σ_b . In the optimal projector selection, the major variance σ_a is needed in Eq. (12), and is therefore calculated by Eq. (7) and (9):

$$\sigma_a = \frac{\sigma_b}{\cos \alpha}. \quad (13)$$

In the following experiment, we computed the major variance of the PSF from the measured and interpolated/extrapolated data shown in Fig. 6(b) and Eq. (13). Note that the PSF calibration has to be performed for every projector in advance.

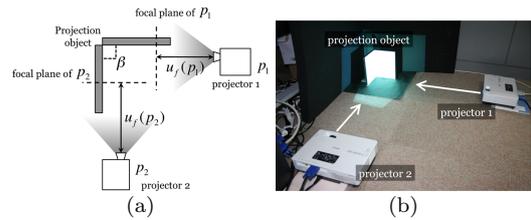


Fig. 7 Experimental setup. (a) Diagram. (b) Actual configuration.

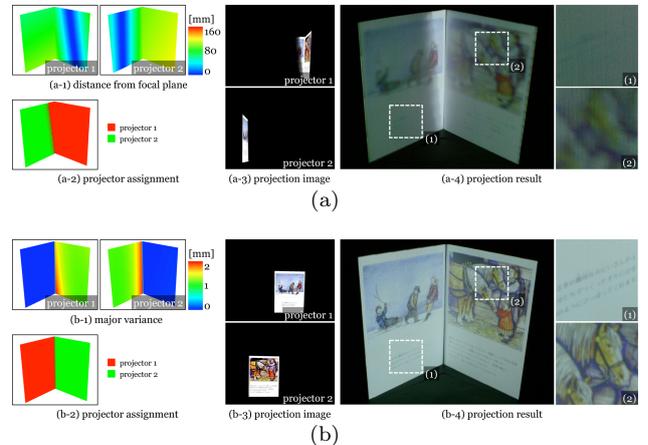


Fig. 8 Experimental result. (a) Naïve approach: the oblique blur was not considered. (b) The proposed approach: both the defocus blur and oblique-blur were explicitly dealt with.

4.2 Significance of Oblique Blur in Defocus Compensation

One of the most important contributions of our approach is that it explicitly considers the oblique blur, in addition to the defocus blur. Thus, we conducted an experiment to validate its effectiveness by comparing our approach with another that deals with defocus blur but not oblique blur (we refer to this as naïve approach). For each point on the surface, naïve approach selects the optimal projector so that its focal plane is the nearest to the point.

We located two projectors and the projection object, as illustrated in Fig. 7. The object consisted of two planar and uniform white surfaces that were connected so that they shared one edge. The projectors projected images of a picture book. The angle of the surfaces β was set to be 90 deg. Shapes of the surfaces and relative rigid-body transforms of the projectors to the surfaces were calibrated in advance.

Figure 8(a) shows the result of naïve approach. For each point of the projection surface, distances of the point from the focal planes of the projectors were calculated (Fig. 8(a-1)), and a projector whose distance was shorter was selected (Fig. 8(a-2)). The images that the two projectors displayed are shown in Fig. 8(a-3).

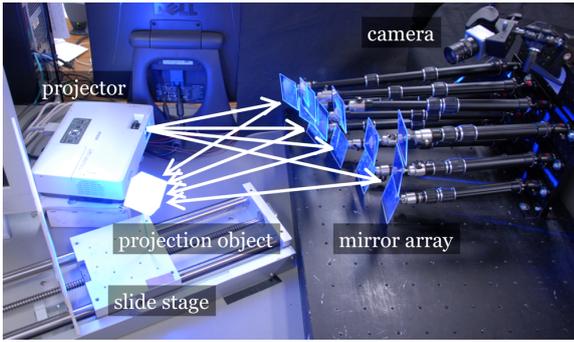


Fig. 9 Experimental setup.

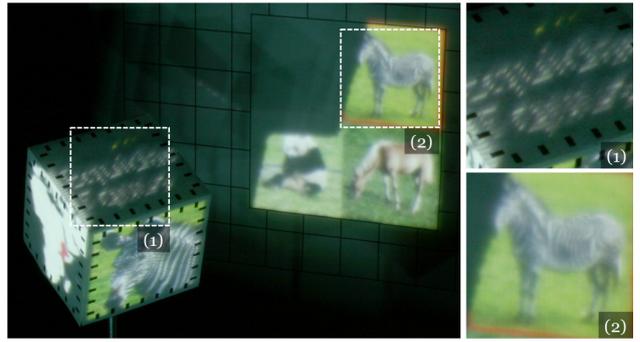
As shown in the projection result (Fig. 8(a-4)), the displayed content was highly blurred.

Figure 8(b) shows the result of our approach. Figure 8(b-1) shows the calculated major variances of the PSFs of both the projectors for each point of the surface. The optimal projectors were selected based on Eq. (12) as shown in Fig. 8(b-2). The projection images are shown in Fig. 8(b-3). Figure 8(b-4) shows the projection result. Comparing Fig. 8(a-4) with Fig. 8(b-4), we could confirm that our approach that explicitly deals with the oblique blur in addition to the defocus blur gave a better result.

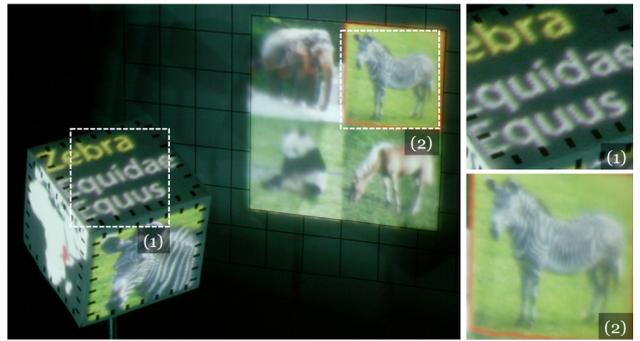
4.3 Dynamic Defocus and Occlusion Compensation

We conducted an experiment to confirm whether our approach could realize dynamic defocus and occlusion compensation. In this experiment, we used a 3×4 mirror array and a projector (parent projector) to practically realize 12 projectors (child projectors) as shown in Fig. 9. The image resolution of each child projector was 225×200 pixels. The similar system configuration was proposed in (Levoy et al, 2004). Unlike the previous configuration where the focal planes of the child projectors were set at the same position, we adjusted the mirrors so that the focal planes of the child projectors differed each other. Two projection objects were located in front of the child projectors. One of the objects was a planar surface, and the other was a cube that could be translated in one direction by a slide stage (Fig. 9). Their shapes were known and represented by 11,000 vertices total.

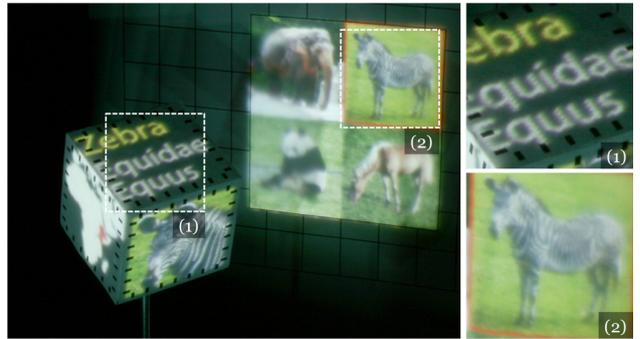
Geometric calibration of the child projectors was performed through a gray-code pattern projection method by using a camera (Point Grey Research, Scorpion, SCOR-20SOM-CS, 1280×960 [pixel]) (Sato and Inokuchi, 1987). The position and orientation of the surface and the movement direction of the slide stage were also measured with the camera (see Appendix A for details).



(a)



(b)



(c)

Fig. 10 Experimental result of defocus-free and occlusion-free projection (left: overall view, right: close-up). (a) Projection result by a single child projector. (b) Projection result by 12 child projectors. (c) Projection result by 12 child projectors after the cube was translated.

Figure 10 shows the projection results of a single child projector (Fig. 10(a)), of 12 child projectors (Fig. 10(b)), and of 12 child projectors after the cube was translated (Fig. 10(c)). Note again that any additional calibrations were not performed in the experiment. In the case of a single child projector, the image quality in the region (1) of Fig. 10(a) was degraded by the oblique blur while a fine image was displayed in the region (2). Furthermore, not all of the image contents were displayed because of the cast shadow. On the other hand, in the case of 12 child projectors, fine images were displayed both in the region (1) and (2), and shadows were successfully removed in Fig. 10(b).

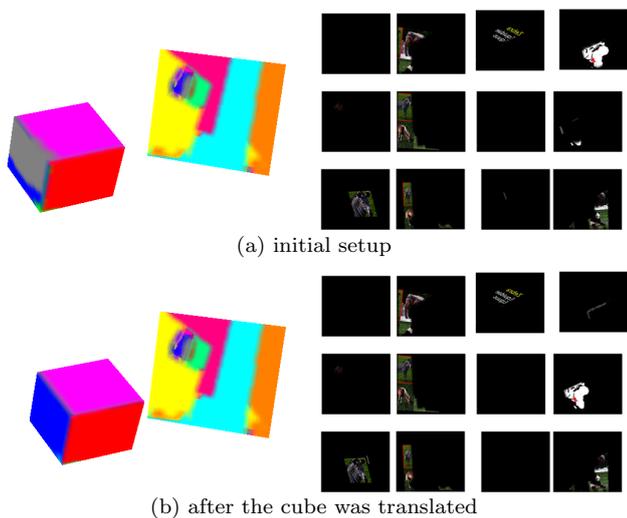


Fig. 11 Child projector contribution. (left) Selected projectors for each point of the projection surfaces (each color corresponds to each child projector). (right) Projection images from the parent projector. Each small rectangle corresponds to each child projector.

When the target moved on the slide stage, defocus and occlusion compensation was also realized as shown in Fig. 10(c). The system selected the optimal projector according to the recalculated major variance of the PSF for each point of the surface and generated the projection image. Consequently, the image quality of the projection result was not degraded. From the result, we confirmed that our approach could keep displaying fine images on the moving surface without any additional calibrations when the 6-DOF movement was known.

Figure 11 shows the selected projectors for every point of the target surfaces and the raw projection image from the parent projector. Figures 11(a) and (b) show the data of the initial setup and after the cube was translated, respectively. We could confirm that the optimal projector selection and the projection image were updated.

5 Discussion

This section briefly discusses advantages and limitations of the proposed method and directions for future work.

5.1 Limitation of Defocus Compensation

One of the limitations of the proposed method is that a defocus free projection cannot always be guaranteed because of the relatively short focal depth of video projectors. Although adding more projectors leads to minimize the problem, in practice there are still some areas

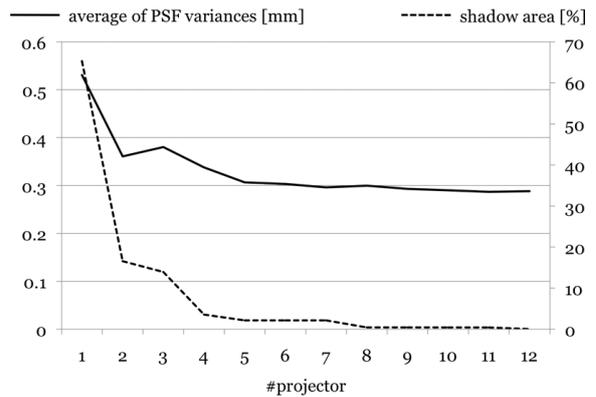


Fig. 12 System performance: average of variances of the PSF of the projected pixels, and percentage of the entire projection area of the shadow area.

where not a single projector is able to project completely in focus. It would occur especially in larger and more complicated scenes than our prototype.

5.2 Photometric Correction

Photometric compensation is needed to correct color differences of results displayed by different projectors. Once initial photometric calibrations are performed, projectors' inherent color properties such as brightness, gamma, and white balance can be compensated through an open-loop process. However, intensity differences occur when the incidence angles or the distances of the projectors to the projection surface are different. Because these are dynamically changed, a closed-loop process is needed to correct projection images. We can apply the following optical models to solve this problem. First, the intensity of the reflection of the projection light is inversely proportional to the square of the distance from the projector to the surface. Second, suppose the surface exhibits Lambertian reflectance; then radiant intensity observed from the surface is directly proportional to the cosine of the incidence angle. Because we already have geometric information of the projectors and the surface, we can calculate the final intensity values displayed on the surface by these models before projection. Thus, we can perform photometric correction to compensate intensity variations among projectors.

5.3 Performance

Although we built 12 child projectors with a mirror array and a parent projector in the experiment of Section 4.3, 12 is not an inevitable number. As shown in Fig. 12, the displayed image's quality was not greatly improved with more than 6 projectors. Therefore, 5 projectors are

sufficient in this case. Further projectors would simply increase the cost and the processing time. Thus, one of our future works would be to develop an algorithm that automatically calculates the minimum number and optimal positions of projectors based on information about the projection object such as the number, shape, and range of movement so that the worst image quality that is specified in advance is guaranteed.

5.4 Optimal Projector Selection Criteria

The previous defocus compensation method compares the areas of the measured PSFs for the optimal projector selection (Bimber and Emmerling, 2006). On the other hand, our method selects a projector whose major variance of the PSF is minimum as the optimal one. We evaluated both the criteria through a psychophysical study.

In the study, participants observed blurred images synthesized with different PSFs and compared the perceived image qualities. We prepared ten different original images (1 English text and 9 color natural images, 512×512 pixels) and three PSFs ((a1) 10×10 , (a2) 16×16 , and (b) 5×18 pixels). For each original image, three blurred images were synthesized with the PSFs (Fig. 13). Consequently, 30 blurred images were synthesized. Each participant observed a set of two synthesized images ((a1) and (b), or (a2) and (b)) and their original image on an LCD monitor. Then he (or she) judged which synthesized image has a better image quality. Each participant performed 20 trials (2 sets, 10 original images). 10 participants (7M, 3F) took part in the study. All participants were naïve to the purpose of the experiment and had normal or corrected to normal vision. As a result, the participants judged that the image quality of (a1) was better than (b) in 89% of all the trials and the image quality of (b) was better than (a2) in 89% of all the trials.

From the result, it is confirmed that when there are two projectors whose PSFs are (a1) and (b) respectively, the projector whose PSF is (a1) should be selected as the optimal projector. On the other hand, when there are two projectors whose PSFs are (a2) and (b) respectively, the projector whose PSF is (b) should be selected. In the first case, our method could select the correct projector because it compares the major variance (i.e., the length of the major axis) of the PSF. On the other hand, the previous method could not select the correct one because the area of the PSF (a1) is larger than that of (b). However our method could not select the correct one in the second case while the previous method could.

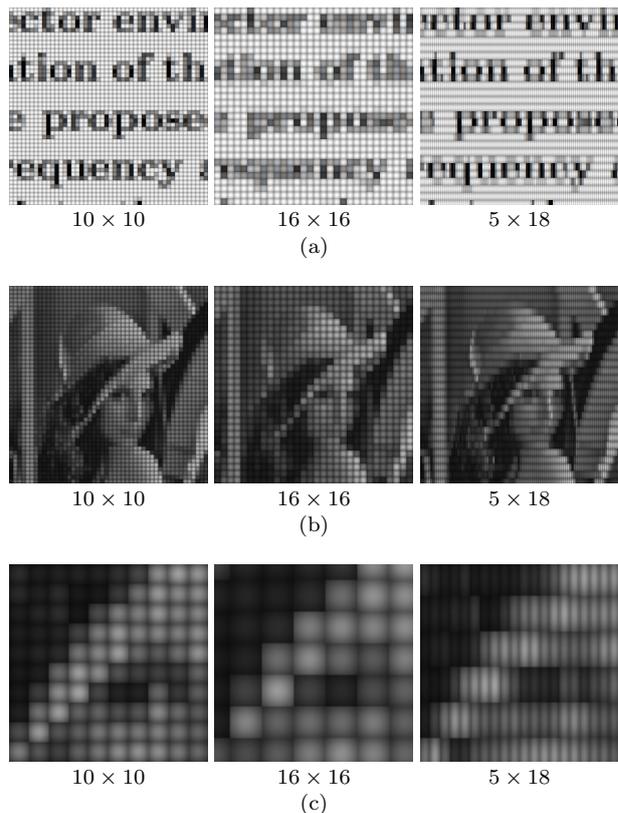


Fig. 13 Example of stimuli. (a) English text. (b) Lenna (a natural color image). (c) Magnified view of Lenna's right eye.

It was indicated that when the areas of the PSFs greatly differ, it would be better to compare the areas for optimal projector selection. On the other hand, when the difference is small, the major variances of the PSFs should be compared. This means that not a single criterion (either major variance or area) but the both are needed in the optimal projector selection. Our model-based method can compute and compare both the criteria, because it computes the shape of the PSF by taking into account the oblique blur. On the other hand, the previous appearance-based (or camera-based) approach cannot compute the exact shape of the PSF, so only the area of the PSF can be compared.

The best solution would be to select the optimal criterion according to the shapes of the compared PSFs. One of our future works would be to investigate the optimal projector selection method which considers both the major variance and the area of the PSF.

6 Conclusion

We presented a novel dynamic defocus and occlusion compensation approach in a multi-projector environment. Our approach compensates for defocus blur, oblique blur, and cast shadows of the projected images based

on geometric information on projectors and projection surfaces as well as frequency analysis of the PSFs. Our approach achieves online and continuous image corrections while the target object moves. In the first proof-of-concept experiment, we confirmed that our approach that considers the both defocus and oblique blurs could display images with a much better quality than another naïve approach that considers only defocus blur to select the optimal projector. The result of the second experiment showed that the proposed approach could realize defocus and occlusion compensation simultaneously for a dynamic scene.

In this paper, we assume that all the projectors are static, or have no pan-tilt-zoom (PTZ) mechanisms. A user can adjust the projecting direction and spatial resolution of the projected image with PTZ projectors. Therefore, as a future work, we will investigate the potentials and possibilities of dynamic occlusion- and defocus-free projection by multiple PTZ projectors.

References

- Audet S, Cooperstock J (2007) Shadow Removal in Front Projection Environments Using Object Tracking. In: Proceedings of IEEE CVPR '07, pp 1–8
- Bandyopadhyay D, Raskar R, Fuchs H (2001) Dynamic Shader Lamps: Painting on Movable Objects. In: Proceedings of IEEE/ACM ISAR '01, pp 207–216
- Bimber O, Emmerling A (2006) Multifocal Projection: A Multiprojector Technique for Increasing Focal Depth. *IEEE Transactions on Visualization and Computer Graphics* 12(4):658–667
- Bimber O, Raskar R (2005) Spatial Augmented Reality: Merging Real and Virtual Worlds. A. K. Peters Ltd.
- Bimber O, Iwai D, Wetzstein G, Grundhöfer A (2008) The Visual Computing of Projector-Camera Systems. *Computer Graphics Forum* 27(8):2219–2254
- Brown MS, Song P, Cham TJ (2006) Image Pre-Conditioning for Out-of-Focus Projector Blur. In: Proceedings of IEEE CVPR '06, vol II, pp 1956–1963
- Cham TJ, Rehg JM, Sukthankar R, Sukthankar G (2003) Shadow Elimination and Occluder Light Suppression for Multi-Projector Displays. In: Proceedings of IEEE CVPR '03, pp 513–520
- Grosse M, Bimber O (2008) Coded Aperture Projection. In: Proceedings of ACM EDT-IPT '08, pp 13:1–13:4
- Gupta S, Jaynes C (2006) The Universal media book: tracking and augmenting moving surface with projected information. In: Proceedings of IEEE/ACM ISMAR '06, pp 177–180
- Hartley RI, Zisserman A (2004) Multiple View Geometry in Computer Vision. Cambridge University Press
- Jaynes C, Webb S, Steele RM (2004) Camera-Based Detection and Removal of Shadows from Interactive Multiprojector Displays. *IEEE Transactions on Visualization and Computer Graphics* 10(3):290–301
- Kondo D, Kijima R (2002) Proposal of a Free Form Projection Display Using the Principle of Duality Rendering. In: Proceedings of VSMM '02, pp 346–352
- Kondo D, Shiwaku Y, Kijima R (2008) Free Form Projection Display and Application. In: Proceedings of IEEE/ACM PROCAMS '08, pp 31–32
- Levoy M, Chen B, Vaish V, Horowitz M, McDowall I, Bolas M (2004) Synthetic Aperture Confocal Imaging. *ACM Transactions on Graphics* 23(3):825–834
- Low KL, Welch G, Lastra A, Fuchs H (2001) Life-Sized Projector-Based Dioramas. In: Proceedings of ACM VRST '01, pp 93–101
- Majumder A, Brown MS (2007) Practical Multiprojector Display Design. A. K. Peters
- Oyamada Y, Saito H (2008) Defocus Blur Correcting Projector-Camera System. *Lecture Notes in Computer Science* 5259:453–464
- Park H, Lee MH, Kim SJ, Park JI (2006) Surface-Independent Direct-Projected Augmented Reality. In: Proceedings of ACCV '06, vol 2, pp 892–901
- Raskar R, Welch G, Low KL, Bandyopadhyay D (2001) Shader Lamps: Animating Real Objects With Image-Based Illumination. In: Proceedings of Eurographics EGWR '01, pp 89–102
- Sato K, Inokuchi S (1987) Range-Imaging System Utilizing Nematic Liquid Crystal Mask. In: Proceedings of IEEE ICCV '87, pp 657–661
- Sukthankar R, Cham TJ, Sukthankar G (2001) Dynamic Shadow Elimination for Multi-Projector Displays. In: Proceedings of IEEE CVPR '01, vol II, pp 151–157
- Zhang L, Nayar S (2006) Projection Defocus Analysis for Scene Capture and Image Display. *ACM Transactions on Graphics* 25(3):907–915

A Geometric Calibration of Projector

We apply a geometric calibration method based on gray-code projection, which was proposed by Sato and Inokuchi (Sato and Inokuchi, 1987).

Suppose that a 3D point (X, Y, Z) in a world coordinate system is projected onto a 2D image plane (x, y) . According to the pinhole camera model, the projection can be described by the perspective equation with the 3×4 perspective projection matrix \mathbf{C} :

$$h \begin{bmatrix} x & y & 1 \end{bmatrix}^t = \mathbf{C} \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^t, \quad (14)$$

where

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & 1 \end{bmatrix}. \quad (15)$$

\mathbf{C} is determined up to a scale factor h , and has eleven unknown parameters. Because two equations are derived from Eq. (14), the unknown parameters of \mathbf{C} can be solved using a least-squares method with six or more correspondences between the 3D world and the 2D screen coordinate systems. The perspective projection matrix can be decomposed to intrinsic and extrinsic matrices (Hartley and Zisserman, 2004). Once the extrinsic matrix is calculated, a 3D rigid-body transform of the pinhole camera device (camera and projector in our case) in the world coordinate system can be computed. The basic idea of geometric registration of each child projector to target objects is to use a camera to determine the relationship (i.e., perspective projection matrix or extrinsic matrix) between the object and the projector.

In the actual calibration process, we position a reference cube (50 [mm] on a side) with spatially-known feature points on the slide stage within the intersection of view frusta of the projector and the camera. The fiducial cube determines the world coordinate system (Fig. 14(a)). First, the camera captures the fiducial cube, and the fiducial points are automatically extracted in the captured image. Then, intersection points of two grid line segments each of which connects two fiducial points are calculated as shown in Fig. 14(b). The number of the intersection points is 147 in total. With 147 correspondences between the world coordinate value (X, Y, Z) and the camera screen coordinate value (x_c, y_c) , a perspective projection matrix \mathbf{C} is calculated using a least-squares method.

$$h \begin{bmatrix} x_c & y_c & 1 \end{bmatrix}^t = \mathbf{C} \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^t. \quad (16)$$

In the geometric calibration of each child projector, horizontal and vertical gray-code patterns are projected onto the fiducial cube from each child projector to compute the perspective

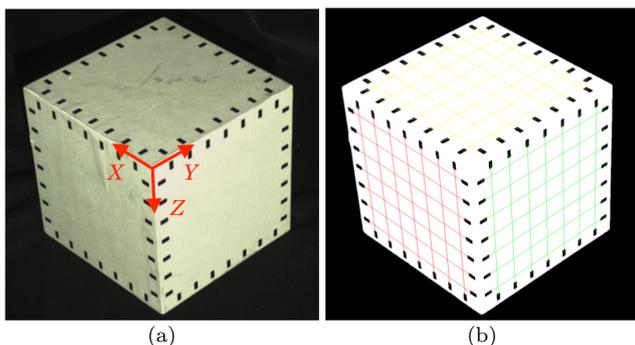


Fig. 14 Fiducial cube. (a) Spatially-known feature points are printed on it. (b) Grid line segments each of which connects detected two fiducial points.

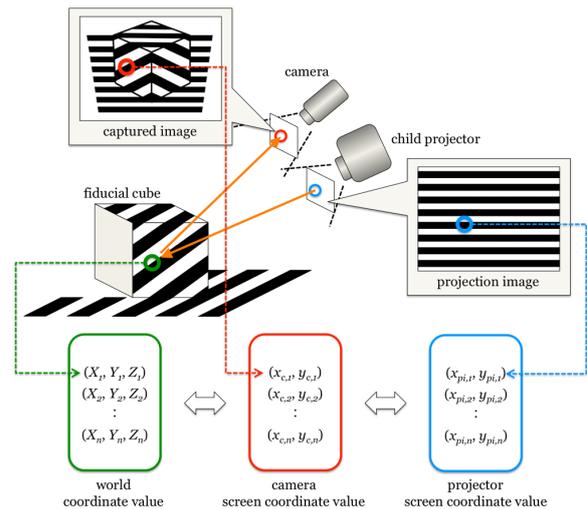


Fig. 15 Gray-code projection to obtain geometric correspondences between the world and the projector screen coordinate systems via camera screen coordinate system.

projection matrix \mathbf{P}_i ($i = 1, \dots, 12$). The projected scenes are captured by the camera (Fig. 15). The captured images are processed and correspondences between camera screen and projector screen coordinate values ($(x_c, y_c) \leftrightarrow (x_{pi}, y_{pi})$) are obtained. Correspondences between world and camera screen coordinate values ($(X, Y, Z) \leftrightarrow (x_c, y_c)$) were already obtained in the camera calibration. Therefore, correspondences between world and projector screen coordinate values ($(X, Y, Z) \leftrightarrow (x_{pi}, y_{pi})$) are derived from these two correspondences (Fig. 15). With the 147 correspondences, the perspective projection matrix of each child projector \mathbf{P}_i is calculated.

$$h \begin{bmatrix} x_{pi} & y_{pi} & 1 \end{bmatrix}^t = \mathbf{P}_i \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^t. \quad (17)$$

Because only a part of child projectors is focused on the fiducial cube, we translate it by using the slide stage so that other child projectors focus on the cube. Then the camera captures the cube again, and the perspective projection matrix \mathbf{C}' is computed. Comparing the extrinsic matrices of \mathbf{C} and \mathbf{C}' , we can compute the rigid-body transform (i.e., translation vector) of the cube. Then we calibrate the uncalibrated child projectors one by one with the same method as mentioned above by taking into account the translation vector of the cube. If there are still defocused child projectors, we move the cube again and repeat the calibration process.

We use the fiducial cube as a projection object. In addition, we place a planar surface behind the cube, on which spatially-known lines are drawn. The system takes a picture of the surface by the camera and recognizes the lines automatically. Because the intrinsic parameter of the camera is derived from \mathbf{C} , we can compute the rigid-body transform of the surface from the camera. Consequently, position and orientation of the surface in the world coordinate system is calculated. When the cube is translated by the slide stage, the system captures it and automatically compute the position and orientation of the cube in the world coordinate system.